sciendo

PROCEEDINGS OF THE LATVIAN ACADEMY OF SCIENCES. Section B,
Vol. 78 (2024), No. 4 (751), pp. 233–238.

DOI: 10.2478/prolas-2024-0032

Review

TOWARDS A HUMAN-AI HYBRID MEDICINE: FUTURE MEDICINE — A HYBRID SYSTEM WHERE AI COMPLEMENTS INSTEAD OF REPLACES HUMANS

Jurģis Šķilters[1], Juris Pokrotnieks[2,3], and Aleksejs Derovs[2,4]

[1] Laboratory for Perceptual and Cognitive Systems, Faculty of Science and Technology, University of Latvia, 19 Raiņa Blvd., Rīga, LV-1586, LATVIA

[2] Department of Internal Diseases, Rīga Stradiņš University, 16 Dzirciema Str., Rīga, LV-1007, LATVIA

[3] Pauls Stradiņš Clinical University Hospital, 13 Pilsoņu Str., Rīga, LV-1002, LATVIA

[4] Rīga East University Hospital, 2 Hipokrāta Str., Rīga, LV-1038, LATVIA

[#] Corresponding author, jurgis.skilters@lu.lv

Communicated by Dainis Krieviņš

Our paper provides a critical overview of the advantages, disadvantages, uncertainties, and challenges regarding AI application in medicine. Without denying the importance of the AI in medical applications, we are arguing for a hybrid and complementary view of future medical systems where powerful AI resources are integrated in and with human decision making.

Keywords: *applications of AI in medicine, artificial and natural cognitive systems, medical safety, medical sustainability.*

INTRODUCTION

Although AI in medicine has tremendous potential and lots of current fruitful uses, it has also substantial problems, limitations, and challenges that we do not know how to overcome yet (Beam *et al*., 2023). And even some of the problems that we are aware of cannot be solved within the existing approaches. WHO calls for key regulatory considerations on AI for health (WHO, 2023), arguing for a dialogue between stakeholders.

The potential of medical AI is critical for contemporary society and its long-term future. The aging society and the current medical establishment might not be able to deal with it without the use of powerful AI systems. Also, the development of new techniques, precision medicine cannot be developed without the use of AI tools.

At the same time, there are several problems and challenges: most notably, almost all medical AI systems rely on large-scale data sets that are used via powerful machine learning techniques. The data-driven approach is the core of contemporary medical AI systems. After a brief description of some limitations in the definition of AI, and misidentification of AI with the data-driven (bottom-up) approach, we provide an overview of crucial benefits and limitations that medical AI systems are having. The shortcomings of data-driven approaches and psychological resistance regarding AI applications are discussed. Finally, a solution of a sustainable medical system is suggested by complementing human decision making and powerful AI tools.

WHAT DO WE MEAN BY AI?

Without attempting to define AI in a way that is comprehensive for 2024 and accurate at the same time, it is worth bearing in mind that AI in recent developments is frequently misidentified as machine learning or bottom-up systems in general. These systems use large data sets as their main input and, depending on the size of the data-sets and the computational power involved, they might achieve powerful and surprising results. Primarily bottom-up principles are the ones that are implemented in ChatGPT type of AI systems, and more broadly machine learning, deep learning, and neu-

ral networks including. At the same time, there are also logic-based or top-down systems that are historically prior to the machine-learning type of AI systems. And even if the bottom-up systems are considered, their scope is much wider than just dialogue systems (as in the case of ChatGPT).

In general, the development of current powerful autonomous AI systems is a result of a technological evolution (particularly, bottom-up systems). It is nothing sudden, unexpected, or surprising if considered within the larger systematic and historic context of AI research and development.

Also, the visions underlying the current AI research are much narrower than AI might actually equip us in future and currently contain two large types (Messeri and Crockett, 2024): (a) predictive AI systems — based on machine learning applied on the given data-sets they provide the users about most probable predictions, forecasts, taxonomies, and classifications, and (b) generative AI systems (again in bottom-up, data driven way) mimicking and generating processes similar to human cognition (e.g., language processing in dialogue systems and decision-making embedded in these systems). In practical terms, there are lots of tasks in science that these two types of visions can and do already now to some extent cover (besides just summarising information). However, it should be considered that there are lots of options where AI can be used in future and that do substantially exceed the current visions of AI, e.g., generating causal reasoning systems — a field that the current AI systems are weak at.

Another misunderstanding is an attempt to identify AI with human intelligence. In fact, most of the powerful AI systems operate on different principles than humans do. To generate AI similar to human intelligence might be inefficient. Current AI systems work in a different way (sometimes better, sometimes worse) than humans. In which sense different? Human brain and intelligence operate in an organismic and interactive way and are inherently active whereas artificial intelligence is a passive model (Chemero, 2023; Pezzulo *et al*., 2024). The living organism is anchored to a real-time body and world interactions in the way that it is controlling and reacting to the consequences of action (Pezzulo *et al*., 2024). Humans are active inference systems, and the best models in human intelligence are generated by actively testing them and generating inferences while operating. "Living organisms and active inference agents learn their generative models by engaging in purposive interactions with the environment and by predicting these interactions. This provides them with a core understanding and a sense of mattering, upon which their subsequent knowledge is grounded" (Pezzulo *et al*., 2024).

Considering the application of AI in medical settings there are several advantages and disadvantages and several subtler issues, i.e. drug development, assistance in operating of medical equipment, etc.

ADVANTAGES OF AI IN MEDICAL APPLICATIONS

By using large-data sets, it is possible to achieve insights from unexpected sources and surprising connections in the datasets that humans might miss when observing concrete cases and deriving consequences from a limited number of observations (Rajpurkar *et al*., 2022).

Further, AI can provide support in conditions where clinicians are fatigued. Considering the situation with the aging population, society might have a problem of lacking a sufficient number of medical experts. In this context, the workload of medical experts will most likely increase in the future, and medical AI systems might have a crucial value to secure the core functions of medical system. Enhancing images, measuring and comparing abnormalities, and red-flagging the test results might be of critical importance, especially in diagnostic and therapeutic scenarios. In a scaffolding scenario, the final decision would be in the hands of a medical expert. Nevertheless, one of the most promising implications of AI, where we are anticipating practical success, is in drug design and development. AI could potentially help us design drugs with desired structures and physicochemical properties, thereby aiding in the achievement of tailored therapies.

However, independent of future societal needs constrained by changes in the demographic composition of the society, AI systems can strengthen clinical trials, improve diagnostics (recognition of symptoms etc.), treatment, and possible development of further dynamics based on existing conditions (e.g., by running simulations according to the set of variables in each concrete case) (WHO, 2023). Additionally, the complexity of contemporary medicine exceeds the individual expert knowledge and performance capacity (Goldberg *et al*., 2024).

PROBLEMS AND NEGLECTED ISSUES IN MEDICAL AI SYSTEMS

Firstly, a large set of problems are linked to the so-called AI paternalism. In medicine paternalism relies on the "doctor knows best"-principle. An "All knowing doctor" is a socio-historically established role with lots of psychological and ethical issues (for some of the discussions see Veatch (2000) and Taylor (2009).

However, relying on AI as an "all knowing doctor" is even worse, since faith in technology is highly unreliable because medical technology is dependent on industrial world and interests. Therefore, algorithmic paternalism in medicine is dangerous (Hamzelou, 2023), both because of industrial interests that might be discrepant to the view of a sustainable development in medicine, but also because of the technological constraints of the datasets used in AI systems.

Further, what is good in terms of Health-AI for clinicians is different from advantages for patients (and the other way around). Different perspectives in the applications of AI should be tested in a more depth longitudinally both in

terms of their clinical outcomes but also psychological consequences for the different user groups.

Secondly, AI is in principle not infallible because it is dependent on large human-generated data-sets and the algorithms underlying them. Further, AI cannot *understand* (because understanding requires resonating with developmentally generated sensory and bodily experience and social experience) but can only *predict* to a certain degree according to the input data set. It is worth keeping in mind that *pattern recognition* and its *interpretation* are also different processes. Therefore, AI power in pattern recognition is not necessarily the power in the care of patients (Lenharo, 2023). Most of the current medical AI systems are designed on relatively narrow domains and lots of potentially important variables are still not included yet. A dangerous prospect here is a science where more is produced but less understood (Messeri and Crockett, 2024).

The dependency of AI on large data-sets is linked to the problem that these data sets might be biased according to their sources and the way they are collected. For more on large data-sets and their issues see the section "Large data sets in medical AI: challenges and issues".

Thirdly, AI systems are industrially generated, which means they are in control of particular companies; once we are dependent on AI-tech, what will or might happen if these companies change their policies / manipulate their (our?) data? And what if they go bankrupt? The moral of this is that the development of powerful AI systems is financially motivated (see Hamzelou, 2023).

Loss of control as a general consequence of the existing large-scale autonomous AI systems was a warning emphasised in a position paper by a group of leading intellectuals, including founders of modern AI technologies (Yoshua Bengio and Geoffrey Hinton) and Nobel Prize Winner Daniel Kahneman (Bengio *et al*., 2023).

A large number of problems arise from the fact the there are no clear regulations between medical institutions and industrial partners and there are no clear commitments from industry regarding the use of the data-sets in the way that it is in the interests of society and not just industry; this would mean to take into account also the underserved populations that are not able to afford these tools (see Goldberg *et al*., 2024). In addition, one of the challenging issues is that, in clinical settings, we often use clinical practice guidelines to establish minimum diagnostic, treatment, and other criteria. However, it is unlikely that we will be able to unify these tools in the near future, as AI potential is not yet available in the majority of clinics, even in the Western world.

Fourth, cybersecurity threads, avoiding possible misuse and requirements for transparency and documentation of AI systems is a large group of problems lacking a clear solution yet (WHO, 2023; Goldberg *et al*., 2024). Frequently, we do not know the details about the systems of algorithms that are used in Health-AI.

The uncertainty regarding the responsibility of both (a) sources of the data sets used in machine learning systems for AI purposes and (b) the consequences is a crucial problem without a clear solution yet (Rajpurkar *et al*., 2022).

These problems are also related to legal responsibility: who if an AI system entirely takes over medical interaction is responsible about the results? And even if AI is involved, is the legal responsibility distributed? (see Goldberg *et al*., 2024).

Generative AI systems can generate fake-clinical trial data sets to support unverified claims (Naddaf, 2023). A fake and simulated data set can be easily created without any original data collection at all to support a claim. This kind of data fabrication seems to be very realistic and truthful and is difficult to discover. This can potentially lead to extremely dangerous results because fake measurements on non-existent patients can be easily generated and only good trained experts, forensic analysts or biostatisticians can discover these flaws (Naddaf, 2023).

The general problem is that AI systems tend to reduce costs and most likely will be implemented more and more without understanding in-depth how and what they are operating upon and what are the possible and real consequences in the short and long-term future (Hamzelou, 2023). This is an even more radical problem since more than 500 AI-based tools have already been authorised by US food and drug administration (FDA) for medical applications (Lenharo, 2023).

RESISTANCE TO AI IN MEDICINE

The reception of AI seems to be polarised: on the one hand, a hyper-enthusiastic, and uncritical and, on the other hand, a critical, hesitant, and resistant approach. The problems regarding the uncritical approach were covered in the previous sections. The reason for hesitancy in respect to AI arises from several somewhat subtle factors. The most crucial seems to be a bias entitled 'Uniqueness neglect': "*Uniqueness neglect,* a concern that AI providers are less able than human providers to account for consumers' unique characteristics and circumstances" (Longoni *et al*., 2019). The individuals who consider themselves as more unique tend to be more reluctant in respect to AI. Uniqueness neglect can be reduced by (a) providing more personalised AI systems, (b) acceptance among more individuals within the same society, and (c) supporting instead of replacing final human decisions (Longoni *et al*., 2019).

LARGE DATA SETS IN MEDICAL AI: CHALLENGES AND ISSUES

Large data sets (that are used in machine learning) provide good results for dealing with relatively frequent results, but not that good and reliable when dealing with infrequent/atypical results. This is one of the reasons why medical AI should be considered as a scaffolding system instead of an

autonomous one providing final and conclusive decisions in clinical settings.

Further, data-sets that are used in medical AI systems are frequently skewed and biased. Where do they come from? These are historic data-sets, specific to particular populations and particular periods. These data-sets do not cover the demographic diversity in terms of race, age, class, ethnicity etc. (Messeri and Crockett, 2024). Our findings are highly limited and constrained once we generalise results from particular populations (because of interrelated geographic, biological, and demographic consequences) and times (the symptom strength and the underlying causal pattern changes during time).

Medical AI tools are sensitive according to the relative frequency of the phenomena covered by the data sets and are frequently skipping individual differences that might also be important.

Finally, and especially if the medical AI is considered to be similar to human intelligence involved in health care: Are the tools based on these large data-sets accurate in terms of simulating human cognitive processes if canonical experiments from cognitive psychology are applied on them? Indeed, lots of classical cognitive processes (e.g., vignette-based tasks, decision making from descriptions) can be simulated in a way that is similar to humans; however generative AI systems (like ChatGPT) are relatively weak at causal reasoning (Binz and Schulz, 2023), which is central in medical reasoning. This is another reason why medical AI should be considered as a part of a complementary medical system with human reasoning and decision making.

A large group of problems is related to the issue about models. What are the models that AI experts talk about? Usually they are bottom-up, data-driven and because of that skewed in temporal and population sense. These issues can, of course, be improved but there does not seem to be an inherent solution.

However, there is another way of using the concept 'model' that is established in mathematics — model theory — this is something the current approaches in AI cannot deal with. This is the sense in which a model is the criterion of determining truth about something. This is one of the problems many leading scholars — Geoffrey Hinton, the founder of mainstream neural network technology, Gary Marcus, and others — have emphasised (Bengio *et al*., 2023; Marcus, 2023).

How can we know that that a particular deep-fake is not true? Sophisticated techniques are necessary. But in case of medicine this might be even more problematic and frightening since there are technological and otherwise interests and at the same time human life might depend on these medical AI tools.

Additionally, a widespread (sometimes implicit) motivation is to use the AI to increase productivity, accuracy, and objectivity in science; the problem here is that the AI systems are typically reflecting our own limitations (they have been trained on our data and the standpoints of their developers that are also having a limited and imperfect picture of the world), biases and illusions of understanding (we seem to think that we know more about the world than we actually do) (Messeri and Crockett, 2024).

OTHER PROBLEMS MEDICAL AI IS FACING WITH

Firstly, external validation is a challenge that has not been systematically dealt with: most of the AI tools are not tested on data sets that are entirely independent of the data sets that these tools were trained upon (Lenharo, 2023; for a more general and programmatic critique see Marcus, 2023).

Secondly, prospective evaluation is important but not always done — different metrics should be compared and different sets of variables should be involved. Simulations might be important to see in the long run outcomes when different variables are involved and measured according to the different metrics.

Thirdly, use cases are crucial to see micro-level patterns of causal intervention or diagnosis dynamics. Use cases are not frequently done before medical AI systems are started to be implemented.

Fourthly, most of the medical AI systems lack causal validation and confirmation (Rajpurkar *et al*., 2022; WHO, 2023) which is the consequence of a larger problem: AI tools usually lack experimental testing. Even less often there are longitudinal analyses and validations. This raises questions about the safety and reliability of those medical AI systems.

Sadly, most of the medical AI-hype-proponents have a rough idea about experiments and scientific causality. Relative frequency in large scale data sets (which is what the modern medical AI systems are based upon) is not causal evidence. In this respect it should be kept in mind that medicine is very much different from the majority of other areas where AI is currently applied (Beam *et al*., 2023) — the causal answers are crucial to understand the reasons and effects of medical interventions in the widest sense of the word. In general, to decide whether a medical AI system is plausible, applicable, safe, and providing positive results, experiments are needed. This is very rarely the case with the current AI systems. Most of them lack any experimental evidence.

Fifthly, AI systems suffer from systematic biases because the data are collected from specific populations (e.g., underrepresentation of particular demographic groups), specific time periods, and linked to overemphasis of specific treatments — sometimes it is a virtue of specific data sets that the system is trained upon, but sometimes one cannot exclude commercial interests in emphasising specific treatments and medicals (see also Meskó and Topol, 2023). Additionally, there are purely technical issues that the current AI systems cannot perfectly deal with, e.g., LLMs (Large

Language Models) are sometimes hallucinating (generating outputs without grounding to input data or factual information).

Sixthly, the overall intensity of AI use in medical sciences (and science in general) seems to generate scientific monoculture characterised by specific and narrow type of methods, questions and generating overall less innovative environment, which is primarily driven by the demand of quantitative productivity instead of quality (Messeri and Crockett, 2024).

These and other problems suggest that clinical decision making should not be substituted but at most complemented by medical AI systems (Meskó and Topol, 2023; Goldberg *et al*., 2024).

CONCLUSION: SOLUTIONS AND PROSPECTS

Recent and prominent critiques of AI argue that the uncritical overreliance on AI can result in generation of substantial methodological limitations (causing the scientific monocultures), quantity and speed (at less money) over quality in medical and other research fields. Although it is clear that AI systems will become an integral part of medical research and practice, a more differentiated and critical approach is needed. A prospectively fruitful approach is to use hybrid medical systems, where experts are supported by powerful AI systems. The use of AI will significantly increase and it is worth considering how medical AI systems can increase the quality of health care instead and without relying on them. It is plausible to consider the using of medical AI in the same way the other well-established technologies are used. Hybrid (instead of human decision-making substituting) technologies are the key that is supported by the evolutionary analysis. Humans have been successful in science because of complementing internal (biological) and external (written, digital, computational) tools; science as evolutionary success is the result of human hybrid processing.

To understand the impacts of artificial systems and to detect possible danger we cannot rely on the here-and-now results provided by those systems, but we should take into account the overall dynamics of social and cognitive structures.

Additionally and more importantly for medicine in particular and science in general — hybrid systems would avoid generating (a) illusion of explanatory depth and breadth of AI (illusions about a deeper and wider understanding than a system actually does), (b) narrower set of hypotheses and methods produced by an AI system than might be applicable, (c) illusion of objectivity (according to which the AI is incorporating all possible perspectives and standpoints and therefore should be relied on), and (d) scientific monocultures assuming one approach as the default and suppressing or eliminating other approaches and perspectives and therefore reducing the methodological diversity that is crucial in scientific knowledge production.

Received 5 July 2024
Accepted in the final form 24 July 2024

REFERENCES

Beam, A. L., Drazen, J. M., Kohane, I. S., Leong, T. Y., Manrai, A. K., Rubin, E. J. (2023). Artificial intelligence in medicine. *New England J. Med.*, **388** (13), 1220–1221.

Binz, M., Schulz, E. (2023). Using cognitive psychology to understand GPT-3. *Proc. Nat. Acad. Sci.,* **120** (6), e2218523120.

Bengio, Y., Hinton, G., Yao, A., Song, D., Abbeel, P., Harari, Y. N., Zhang, Y. Q., Xue, L., Shalev-Schwartz, S., Hadfield, G., *et al.* (2023). Managing AI risks in an era of rapid progress. *Science*, **384** (6698), 842–845. DOI: 10.1126/science.adn0.

Chemero, A. (2023). LLMs differ from human cognition because they are not embodied. *Nat. Hum. Behav.*, **7** (11), 1828–1829.

Donald, M. (1993). Précis of origins of the modern mind: Three stages in the evolution of culture and cognition. *Behav. Brain Sci.,* **16** (4), 737–748.

Goldberg, C. B., Adams, L., Blumenthal, D., Brennan, P. F., Brown, N., Butte, A. J., Cheatam, M., deBronkart, D., Dixon, J., Drazen, J., *et al.* (2024). To do no harm — and the most good — with AI in health care. *NEJM AI*, **1** (3), AIp2400036.

Hamzelou, J. (2023). Artificial Intelligence is infiltrating health care. We shouldn't let it make all decisions. *MIT Technol. Rev.,* April 21

Lenharo, M. (2023). An AI revolution is brewing in medicine. What will it look like? *Nature,* **622***,* 686–688.

Longoni, C., Bonezzi, A., Morewedge, C. K. (2019). Resistance to medical artificial intelligence. *J. Consumer Res.*, **46** (4), 629–650.

Marcus, G. (2023). Hoping for the best as AI evolves. *Communi. ACM*, **66** (4), 6–7.

Messeri, L., Crockett, M. J. (2024). Artificial intelligence and illusions of understanding in scientific research. *Nature*, **627** (8002), 49–58.

Naddaf, M. (2023) ChatGPT generates fake data set to support scientific hypothesis. *Nature*, **623** (7989), 895-896.
DOI: https://doi.org/10.1038/d41586-023-03635-w.

Pezzulo, G., Parr, T., Cisek, P., Clark, A., Friston, K. (2024). Generating meaning: Active inference and the scope and limits of passive AI. *Trends Cognit. Sci.*, **28** (2), 97–112. DOI:
https://doi.org/10.1016/j.tics.2023.10.002.

Rajpurkar, P., Chen, E., Banerjee, O., Topol, E. J. (2022). AI in health and medicine. *Nature Med.,* **28** (1), 31–38.

Meskó, B., Topol, E. J. (2023). The imperative for regulatory oversight of large language models (or generative AI) in healthcare. *npj Digital Medicine*, **6** (1), 120.

Taylor, K. (2009). Paternalism, participation and partnership — the evolution of patient centeredness in the consultation. *Patient Educ. Counsel.,* **74** (2), 150–155.

Veatch, R. M. (2000). Doctor does not know best: Why in the new century physicians must stop trying to benefit patients. *J. Med. Philos.,* **25** (6), 701–721.

WHO (2023). WHO Outlines Considerations for Regulation of Artificial Intelligence for Health.
https://www.who.int/news/item/19-10-2023-who-outlines-considerations-for-regulation-of-artificial-intelligence-for-health (accessed 11.08.2024).

CEĻĀ UZ CILVĒKA UN MĀKSLĪGĀ INTELEKTA HIBRĪDMEDICĪNU: NĀKOTNES MEDICĪNA – HIBRĪDA SISTĒMA, KURĀ MĀKSLĪGAIS INTELEKTS PAPILDINA, NEVIS AIZSTĀJ CILVĒKU

Mūsu rakstā sniegts kritisks pārskats par priekšrocībām, trūkumiem, neskaidrībām un izaicinājumiem saistībā ar mākslīgā intelekta (MI) lietošanu medicīnā. Nenoliedzot mākslīgā intelekta nozīmi medicīnā, mēs iestājamies par hibrīdu un papildinošu skatījumu uz nākotnes medicīnas sistēmām, kur spēcīgi MI resursi ir integrēti cilvēku lēmumu pieņemšanā.